



## SHORT RESEARCH COMMUNICATION

# Probability analysis for consecutive-day maximum rainfall for Tiruchirapalli City (south India, Asia)

R. Mani Sabarish<sup>1</sup> · R. Narasimhan<sup>1</sup> · A. R. Chandhru<sup>1</sup> · C. R. Suribabu<sup>1</sup> · J. Sudharsan<sup>2</sup> · S. Nithiyanantham<sup>3</sup>

Received: 27 May 2014 / Accepted: 23 June 2015

© The Author(s) 2015. This article is published with open access at Springerlink.com

**Abstract** In the design of irrigation and other hydraulic structures, evaluating the magnitude of extreme rainfall for a specific probability of occurrence is of much importance. The capacity of such structures is usually designed to cater to the probability of occurrence of extreme rainfall during its lifetime. In this study, an extreme value analysis of rainfall for Tiruchirapalli City in Tamil Nadu was carried out using 100 years of rainfall data. Statistical methods were used in the analysis. The best-fit probability distribution was evaluated for 1, 2, 3, 4 and 5 days of continuous maximum rainfall. The goodness of fit was evaluated using Chi-square test. The results of the goodness-of-fit tests indicate that log-Pearson type III method is the overall best-fit probability distribution for 1-day maximum rainfall and consecutive 2-, 3-, 4-, 5- and 6-day maximum rainfall series of Tiruchirapalli. To be reliable, the forecasted maximum rainfalls for the selected return periods are evaluated in comparison with the results of the plotting position.

**Keywords** Rainfall · Return period · Probability distribution · Chi-square test

## Introduction

Several applications in water resources engineering require appropriate estimate of rainfall depth and its return period from available historic data. Estimation of flood in watersheds, water balance studies, water management studies, rainwater harvesting, detention and retention pond design, evapotranspiration estimation, irrigation planning, etc. are some of the examples where rainfall provides a vital input to design and modeling. Planning and development of water resources at the local or regional level require comprehensive and reliable information of hydrological data of the area under investigation. Prober database is needed to assess the water availability of a region, the absence of which can lead to erroneous planning and design. Long period data can provide reliable water resource assessment. The degree of uncertainty increases if the data length is short. Mathew and Vivekanandan (2009) examined the effect of data length on water resource assessment. The results of the study indicate that the lower the data length, the higher is the likelihood of overestimating water resource availability in the regions. Rainfall at a particular place is also known to be influenced by the results of its local/regional atmospheric and geomorphologic environments.

An important aspect in hydrology is to interpret the future probabilities of occurrence from past records of hydrologic events. Vivekanandan and Mathew (2010) chosen probabilistic modeling to fit six different distributions for annual d-day maximum rainfall of different values of d such as 1, 2 and 3 days for the Devgadhabaria region of

✉ S. Nithiyanantham  
bknithu@gmail.com

C. R. Suribabu  
suribabu@civil.sastra.edu

<sup>1</sup> Centre for Advanced Research in Environment, School of Civil Engineering, Sastra University, Thanjavur 613401, Tamilnadu, India

<sup>2</sup> Department of Civil and Environmental Engineering, SRM University, Kattankualthur 603203, Tamilnadu, India

<sup>3</sup> School of Physical Sciences and Femtotechnology, (Applied Energy Resource/Ultrasonics/Bio-Physics Divisions), SRM University, Kattankualthur 603203, Tamilnadu, India

Gujarat (India). Chi square and Kolmogorov–Smirnov tests are used to judge the applicability of the distributions for modeling of the recorded rainfall data. The standard procedure for estimating the frequency of occurrence of hydrological event is frequency analysis. The objective of frequency analysis of hydrologic data is to relate the magnitude of extreme events to their frequency of occurrence using probability distributions.

The study of extreme rainfall events involves the selection of a sequence of the maximum observations from the respective data series. Goswami et al. (2006) examined the trend of daily high ( $R > 100$  mm) and highest ( $R > 150$  mm) rainfall events over a relatively large region covering 1803 stations for the period 1951–2000. The finding of the study shows that there is a 10 % increase per decade in the level of heavy rainfall event since the early 1950s and more than two times increase in very heavy events. Khan et al. (2007) investigated spatial and temporal variability of daily and weekly precipitation extremes in South America. They have proposed a new measure called the precipitation extremes volatility index to measure the variability of extremes. An analysis of their study indicates the increasing trend of daily maximum rain in the Amazon basin. Guhathakurta et al. (2010) carried out the frequency analysis of rain days, heavy rainfall days and also 1-day extreme rainfall, to observe the impact of climate changes on extreme weather events and flood risks in India. The report shows that the frequency of heavy rainfall events is decreasing in major parts of central and north India, while such events are increasing in Peninsular India and also east and north-east India. The present study aims to evaluate the rainfall magnitude for different return periods and also to ascertain the type of probability distribution that best fits the rainfall.

## Study rainfall station and data

Tiruchirappalli City, known also as Trichy, is an urbanized watershed in the Cauvery River basin. The terrain of the city is flat. The city lies at an altitude of 78 m above sea level and is traversed by the rivers Cauvery and Coleroon. The Coleroon river forms the northern boundary of Trichy. There are few hills located within the city, with the prominent among them being Golden Rock, Rock Fort and the one in Thiruverumbur. During heavy downpour times, the low-lying and improper drainage areas of the city are often subjected to inundation. This happens due not only to blockage of drains, but also to undersized stormwater drains. The study on temporal distribution of rainfall will provide useful information to the city planner. The present study uses 100-year continuous daily rainfall data, obtained

from the Indian Meteorological Department located at Pune.

## Analysis methods

### Probabilistic methods

Probability distributions are widely used in understanding the rainfall pattern and computation of probabilities. In the present study, the probability of exceedance of rainfall  $T = m/(N + 1)$ , where  $m$  is the order or rank and  $N$  is the total number of events. It was computed using the Weibull's plotting position formula and applied to the observed rainfall data. The continuous probability distribution log normal (LN), Pearson type III (P III), log-Pearson type III (LP III) and extreme value type I (EV I) were used to evaluate suitable probability functions.

### Log-normal distribution

A log-normal distribution is a probability distribution of a random variable whose logarithm is normally distributed. The maximum rainfall for a particular return period is calculated using the following equation:

$$X_T = X_{av} + k\sigma, \quad (1)$$

where  $X_{av}$  is the mean values,  $k$  is the frequency factor and

$$\sigma = \left[ \frac{\sum (X_i - X_{avg})^2}{N - 1} \right]^{\frac{1}{2}}, \quad (2)$$

in which  $\sigma$  is the standard deviation and  $N$  is the sample size.

The value of  $k$  is determined considering the coefficient of skewness as zero.

### Pearson type III distribution

In this type of probability distribution, the coefficient of skewness is calculated using the formula given below:

$$C_s = \left( \frac{1}{\sigma^3} \right) \left\{ \frac{N}{(N-1)(N-2)} \right\} \sum (X_i - X_{avg})^3. \quad (3)$$

The frequency factor  $k$  is obtained from the theoretical table available for the Pearson type III distribution with skew coefficient (Patra 2001).

### Log-Pearson distribution

This method is extensively used in hydrologic frequency analysis. The traditional fitting procedure consists of transformation of natural data into logarithms and fitting

logarithmic data to a Pearson type III distribution by the method of moments.

### Extreme value distribution

Gumbel probability distribution is widely used for extreme value analysis of hydrologic and meteorological data such as floods, maximum rainfalls and other events:

$$X_T = X_{av} - \left( \frac{\sqrt{6}}{\pi} \right) \sigma \left[ 0.57721 + \ln \left\{ \ln \frac{T}{T-1} \right\} \right]. \quad (4)$$

The above empirical relation holds good when the record length is 100 years or more (Patra 2001).

### Chow method

Chow (1964) derived the frequency factors ( $k$ ) for log-normal distribution and presented a theoretical table. The value of  $k$  can be obtained using the skewness coefficient ( $C_s$ ) and coefficient of variation ( $C_v$ ). In log-normal distribution, these two parameters are related as

$$C_s = 3C_v + C_v^3, \quad (5)$$

in which the coefficient of variation is obtained as follows:

$$C_v = \frac{\sigma}{X_{av}}. \quad (6)$$

The main difference between the log-normal approach and Chow approach is that calculated  $C_s$  value is adopted in case of Chow method, whereas for the log-normal approach  $C_s$  value is considered as zero.

### Chi-square test

A commonly used test for testing the goodness of fit of empirical data to specific theoretical distribution is the Chi-square test (Haan 1994). The goodness of fit between the observed events and the fitted distribution can be tested. The Chi-square value can be determined for each distribution for a particular return period. A relation between the observed number of occurrence  $O_i$  and expected number of occurrence  $E_i$  can be developed as:

$$\chi^2 = \sum_{i=1}^v \frac{(O_i - E_i)^2}{E_i}. \quad (7)$$

This statistical test judges whether or not a particular distribution adequately describes a set of observations by making a comparison between the actual number of observations and the expected number of observations. The  $\chi^2$  distribution has  $\nu (=N - h - 1)$  degrees of freedom, in which  $N$  is the total number of sample data and  $h$  is the number of parameters used in fitting the proposed distribution. The value of  $\chi^2$  for various degrees of

freedom against distribution percentages is available in the  $\chi^2$  table. In hydrology, 95 % level of confidence is considered as the typical value (Patra 2001). The  $\chi^2$  value corresponding to 98 (i.e.,  $\nu = 100 - 1 - 1$ ) is 124. In general, the probability distribution that provides the least Chi-square value is considered as a best-fitting probability distribution in the recorded range for the given data.

### Analysis of data

The maximum rainfall for 1-, 2-, 3-, 4- and 5-day rainfall days were analyzed as per the procedure detailed above. Figure 1 shows the rainfall depth at different return periods computed using the plotting position method (Suribabu et al. 2015). The 1-day maximum rainfall for the city is obtained as 318.9 mm for the 100-year return period. The maximum rainfall values for 2-, 3-, 4- and 5-day consecutive periods, corresponding to 100 years of return period, are obtained to be 366.6, 368.5, 383.5 and 403.3 mm, respectively. The 1-day maximum rainfall corresponding to the 10-year return period is obtained as 152.4 mm. When the difference in rainfall amount is computed between the 1- and 5-day rainfall amount, the variation in the difference in value is found to be insignificant irrespective of the return period. In case of the 4- and 5-day cases, the significant variation is attributed to increased size of return periods. In most of the return periods, the 4- and 5-day maximum rainfall values are found to be close to each other.

Table 1 shows the maximum rainfall amount for the 10-, 50- and 100-year return periods for 1- to 5-day consecutive rainfalls. The average difference in rainfall amount between the 10- and 50-year return periods is 142 mm, whereas between the 100- and 50-year return periods, its average difference is 40 mm. It is clear from these data that

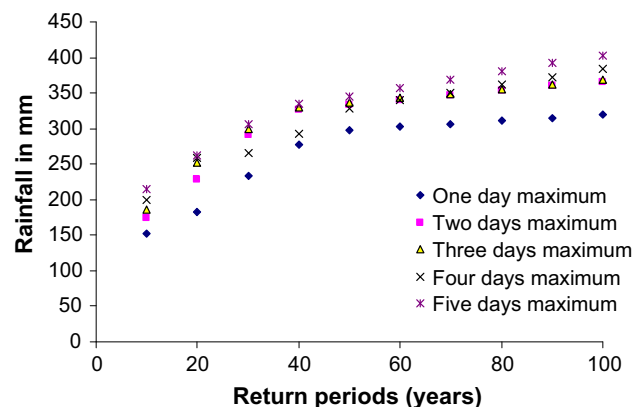


Fig. 1 Maximum rainfall for various return periods (Suribabu et al. 2015)

**Table 1** Maximum rainfall for 1, 2, 3, 4 and 5 consecutive days using the plotting position method (Suribabu et al. 2015)

Return period (year)	One-day rainfall (mm)	Two-day rainfall (mm)	Three-day rainfall (mm)	Four-day rainfall (mm)	Five-day rainfall (mm)
10	152.4	174.8	185.7	199.4	214.4
50	297.96	333	336.2	328.7	345.5
100	318.9	366.2	368.5	383.5	403.3
$X_{50} - X_{10}$	145.56	158.2	150.5	129.3	131.1
$X_{100} - X_{50}$	20.94	33.2	32.2	54.8	57.8

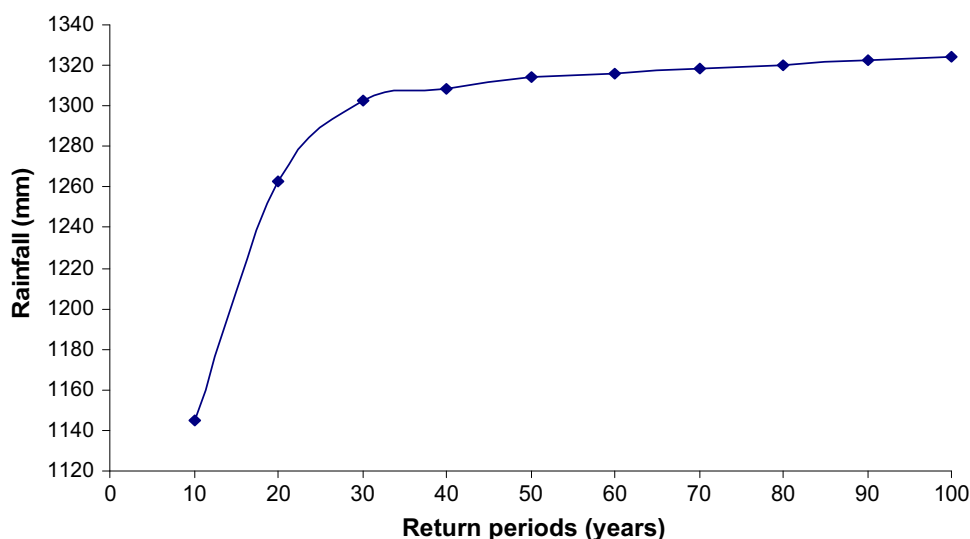
the hydraulic system designed based on the 10-year return period and 50-year return periods can have sizable change in the dimensions of the components. As there is only marginal difference in rainfall amount between the 50- and 100-year return periods, the design based on either 50- or 100-year return period rainfall does not have sizable change in the dimension of the hydraulic components.

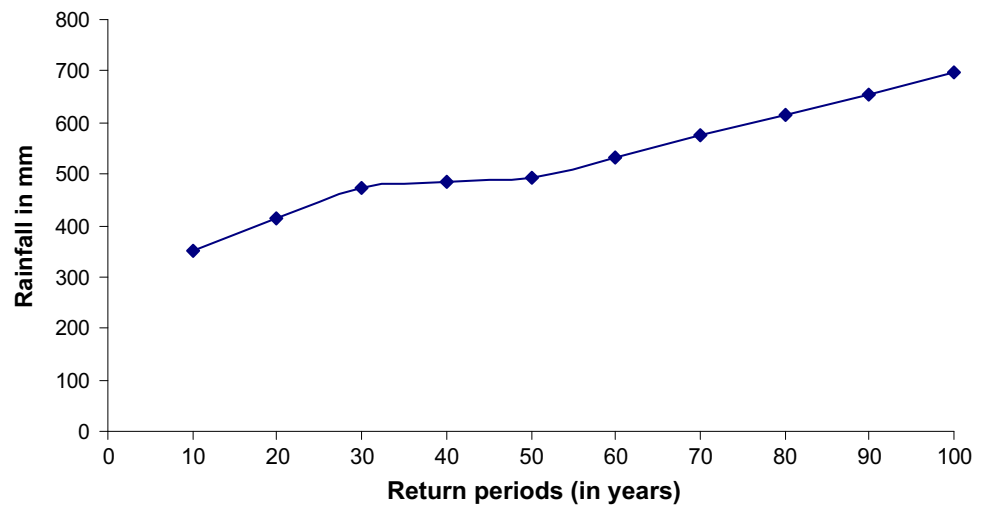
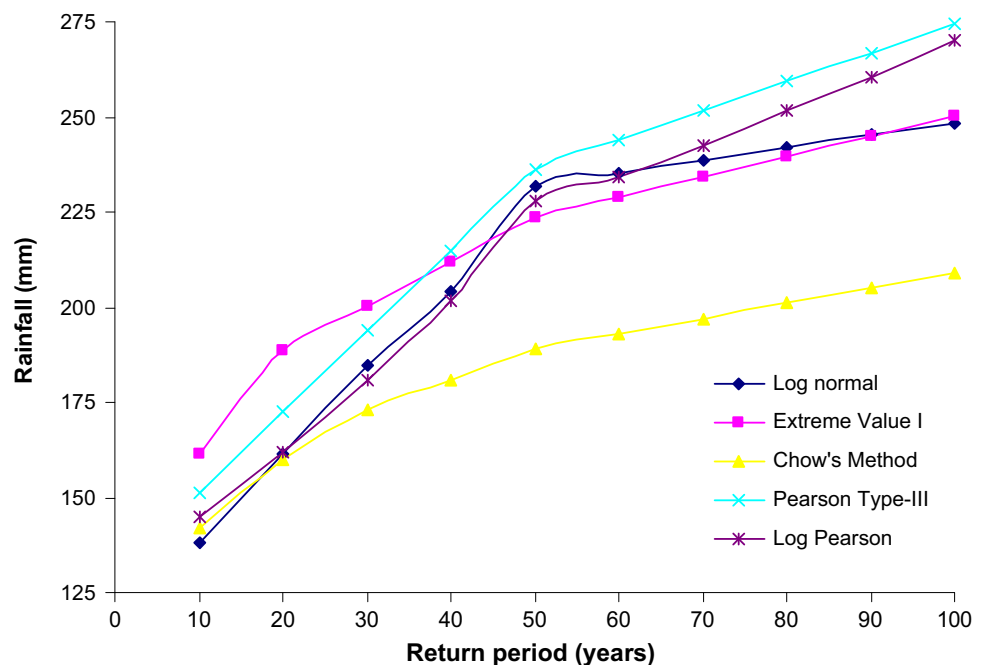
Monthly rainfall for various return periods obtained using the plotting position method (Fig. 2) indicates that the variation of maximum monthly rainfall between the 10- and 50-year return period is obtained as 141.5 mm, whereas that between the 50- and 100-year return period is 204.9 mm. The north-east monsoon period (October–December) provides the maximum monthly rainfall of the study area. From Fig. 2, it may be seen that the slope of curve between 30- and 50-year return period is flatter than other portion of the curve. There is a steep increase in the amount of maximum rainfall from the 50-year return period onward.

While analyzing the annual rainfall for different return periods, the difference in annual rainfall between the 50- and 100-year return periods is 21.7 mm. This indicates that

the 30-year return period is good for planning, in case annual rainfall is required for specific analysis. The annual rainfall figures (Fig. 3) for the 10-, 30-year and 100-year return periods are 1144.6, 1302.8 and 1324.5 mm respectively.

Figure 4 shows the 1 day maximum rainfall for various return periods obtained using different methods. A study of the plot shows that Chow's approach gives the lowest estimate from the 30-year return period onward. Use of the four probability functions LN, PIII, LP III and EV I give results that are close to each other. Among all the methods, the use of P III provides higher estimates. The obtained amount of rainfall for the 100-year return period is 318 mm based on the plotting position method, whereas the value varies from 209 to 275 mm for all the considered probability functions. The highest rainfall amount of 318 mm is recorded between years 1907 and 2006. The value obtained based on Chow's approach gives 209 mm which is 109 mm lesser than the maximum rainfall recorded. The P III distributional model gives 274.5 mm as the maximum daily rainfall for the 100-year return periods which is 43.5 mm less than the maximum recorded value.

**Fig. 2** Maximum monthly rainfall for various return periods

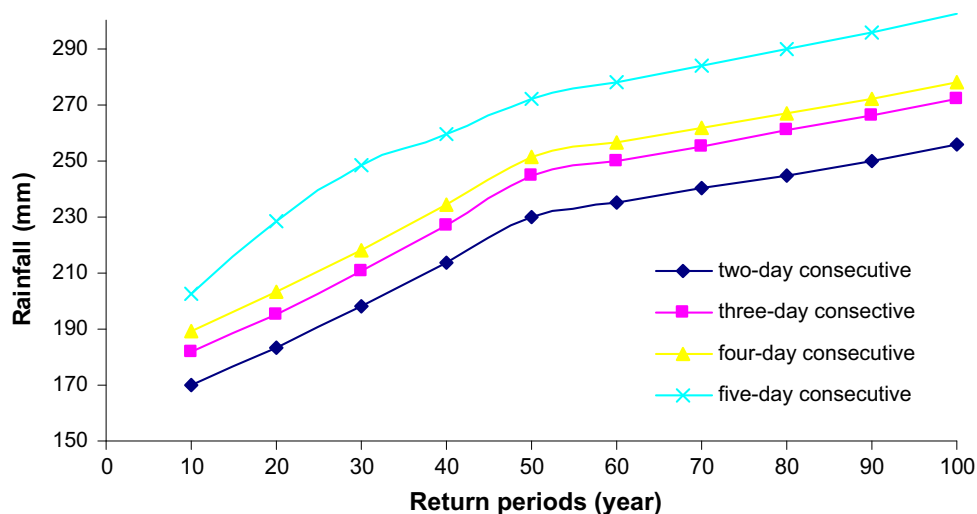
**Fig. 3** Annual rainfall for various return periods**Fig. 4** One-day maximum rainfall based on various probability functions

Figures 5, 6, 7, 8 and 9 show the consecutive-day rainfall at different return periods obtained based on log-normal, extreme value, Chow's method, Pearson type III and log-Pearson, respectively. Upon interpretation of data and further comparison of the curves of the plots prepared between return periods and rainfall, the estimate based on LN distribution shows low values for all consecutive days. The rainfall amount so obtained using other methods provides similar values with  $\pm 10$  mm deviation among each other. These five probability distributions are tested for the goodness of fit by comparing the Chi-square values. The Chi-square value corresponds to the degree of freedom of

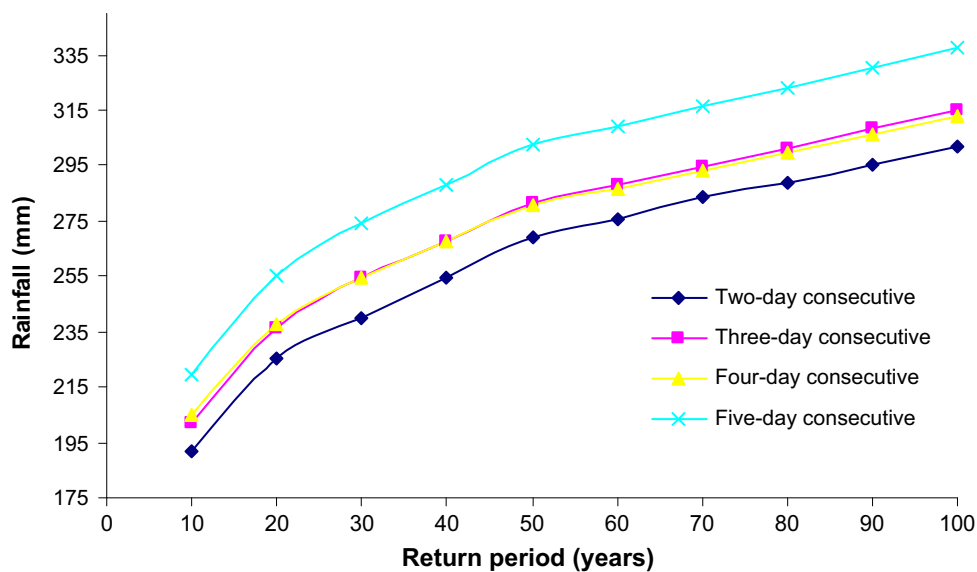
98 %, with the 95 % level of confidence found to be 124 from the Chi-square table available in Patra (2001). While analyzing the Chi-square value for each distribution for the given data, the obtained value of Chi square is less than the table value, hence the hypothesis cannot be rejected at the 95 % confidence level. This indicates that the different models used for fitting the data are accepted.

Figures 10a–e show the Chi-square value obtained for return periods for different probability distribution functions (pdfs). The pdf having least Chi-square value is considered as best-fitting distribution for the given data. The PIII probability function is found to be the best-fit

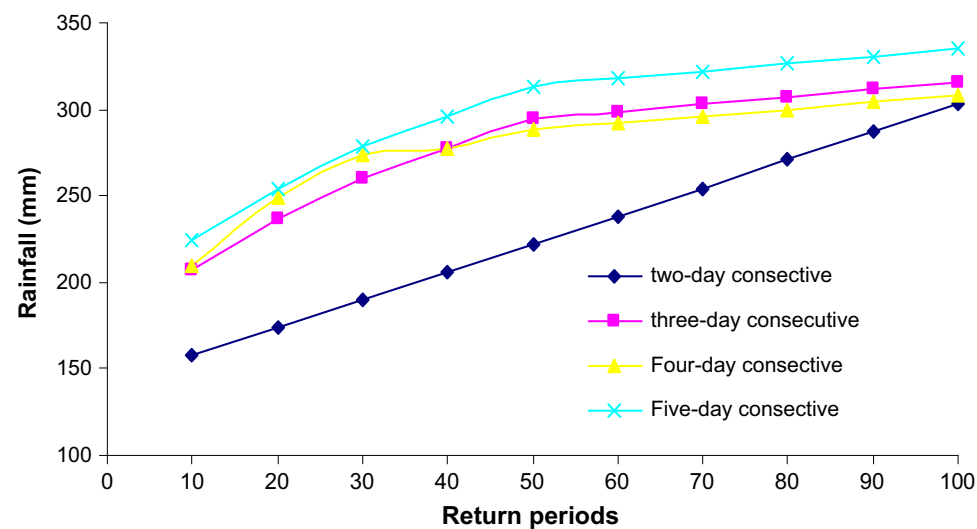
**Fig. 5** Consecutive-day rainfall at different return periods based on log-normal distribution



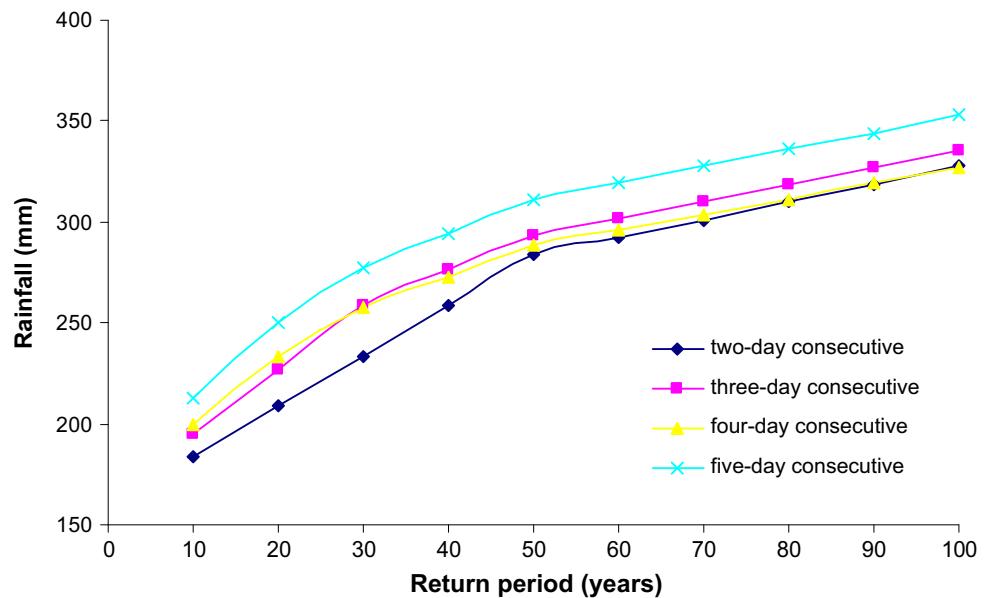
**Fig. 6** Consecutive-day rainfall at different return periods based on extreme value type I distribution



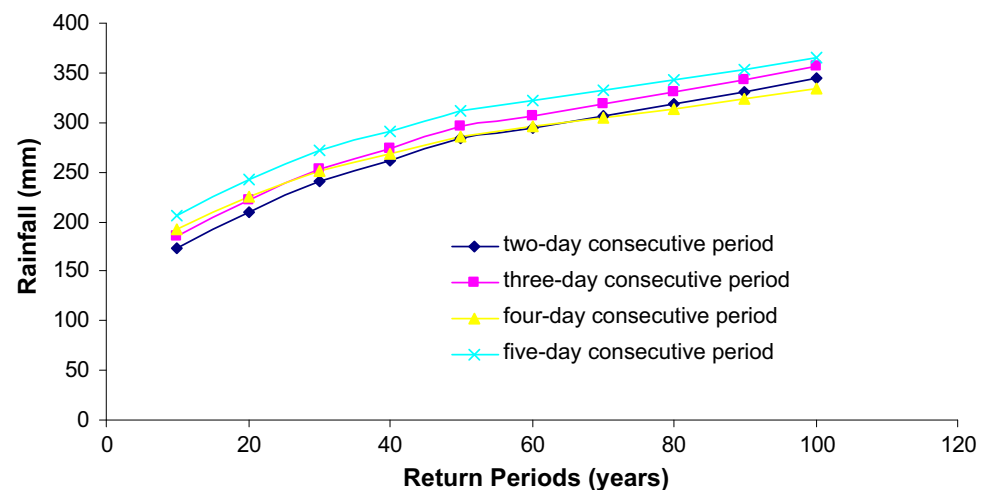
**Fig. 7** Consecutive rainfall at different return periods based on Chow's method



**Fig. 8** Consecutive rainfall at different return periods based on Pearson type III distribution



**Fig. 9** Consecutive rainfall at different return periods based on log-Pearson type III distribution



distribution in case of 1-day maximum rainfall, since it has a minimum Chi-square value for maximum number of return periods. After this, LP III is found to be appropriate and shows the minimum Chi-square values. The curve corresponding to LP III distribution shows low Chi-square values for 2- to 5-day consecutive rainfalls. The difference in rainfall amount between 3- and the 4-day consecutive rainfall is found to be less up to the 60-year return periods and increasing differential amount is observed for the remaining return periods.

For design of drainage system for any urban area, a vital but tricky consideration is the return period of the “extreme” rainfall events. Generally, a best value will lie between overestimating and underestimating the risks involved and a major deciding factor is the cost. When

design is done based on the 10-year return period, the risk involved will be more, whereas if 100-year return period is considered the risk probability will be less. Hence, the selection of appropriate design value is becoming crucial. It is very important that the selection of probability distribution for a particular data set should not provide an underestimated design value. The data corresponding to the 50-year return period can be used in the study area as it falls within the underestimating and overestimating design value. In particular, the design estimate presented here would be beneficial and valuable guidelines during the construction of new drainage systems and rehabilitation of existing drains in the study area, as poor drainage has been identified as one of the major factors causing flooding in the area.

**Fig. 10**  $\chi^2$  Value versus return periods for selected probability functions

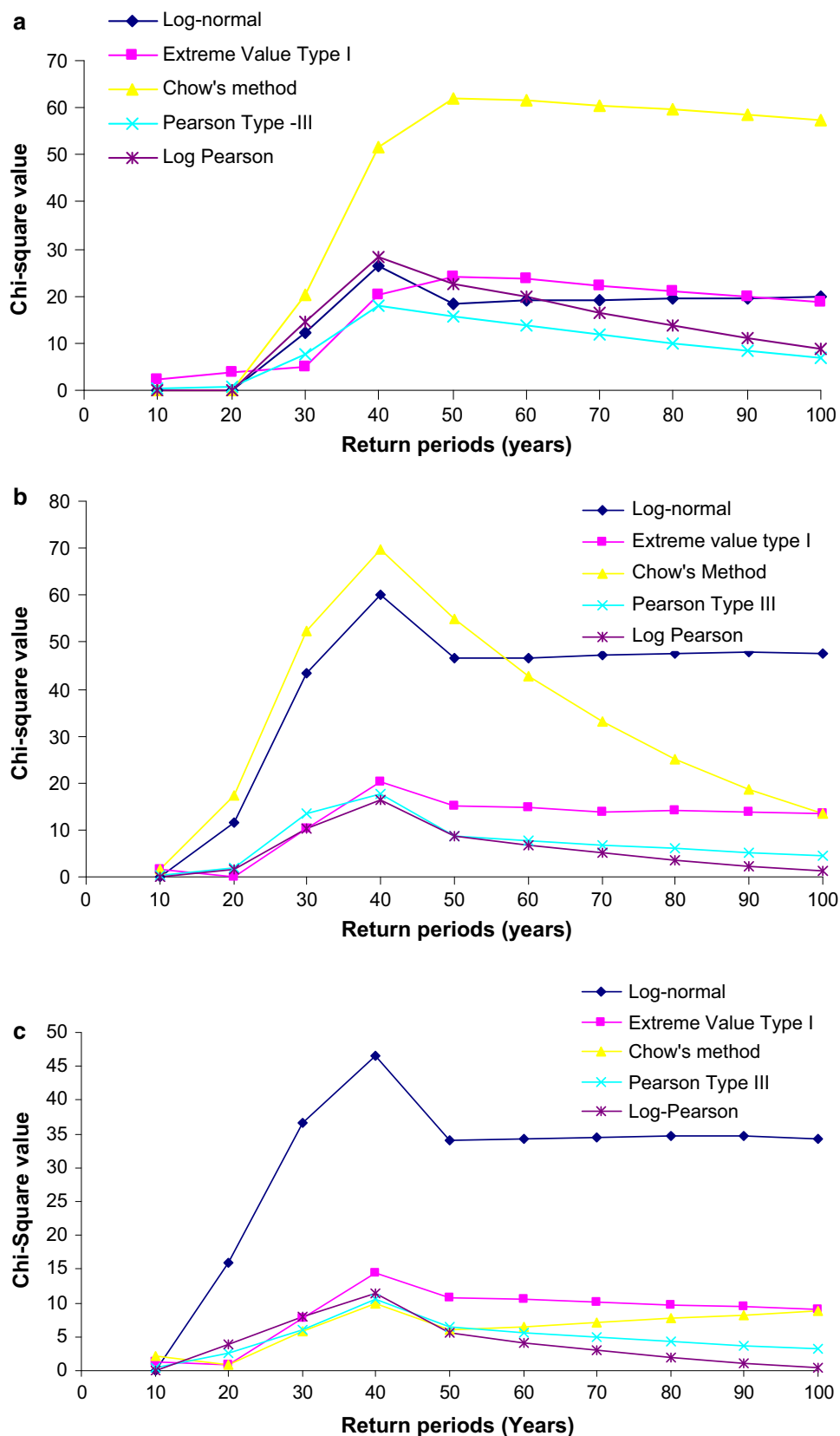
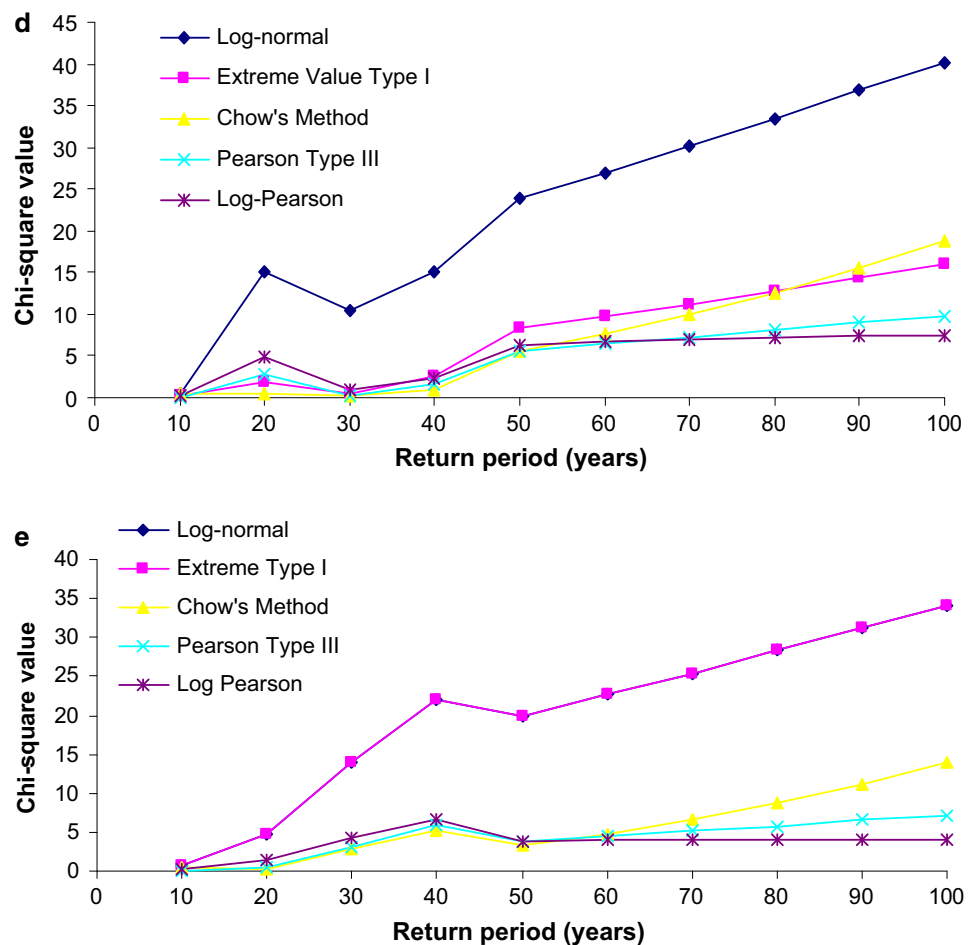




Fig. 10 continued



## Conclusions

Extreme rainfall events for the study area is found out through five different methods and compared with plotting the position method which uses present data by ranking the events as per Weibull's method. The analysis of results indicate that there is a significant difference in rainfall amount between the 10-year and 50-year return periods for 1- to 5-day consecutive rainfall. The probability analysis performed in this study depicts that the difference in rainfall amount for 1- to 5-consecutive day rainfall estimates between the 50- and 100-year return periods is found to be insignificant. Hence, the hydraulic design based on the 50-year return periods holds good even for the 100-year return period rainfall for the study area. In comparison to the estimates based on five probability distribution functions, LP III distribution shows the least value of Chi-square value for the return period up to 100 years and can be adopted for estimation of rainfall amounts at various probability levels.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Chow VT (1964) Handbook of applied hydrology. McGraw Hill book company, New York
- Goswami BN, Venugopal V, Sengupta D, Madhusoodanan MS, Xavier Prince K (2006) Increasing trend of extreme rain events over India in a warming environment. *Science* 314:1442–1445
- Guhathakurta P, Preetha M, Mazumdar AB, Sreejith OP (2010) Change in extreme rainfall events and flood risk in India during last century. National climate centre, Indian meteorological department, Pune, India, Research report no. 3/2010
- Haan CT (1994) Statistical methods in hydrology. Affiliated East West Press Pvt Ltd., New Delhi
- Khan S, Kuhn G, Ganguly AR, Erickson DJ III, Ostrouchov G (2007) Spatio-temporal variability of daily and weekly precipitation extremes in South America. *Water Resour Res* 43(W11424):1–25
- Mathew FT, Vivekanandan N (2009) Effect of data length on water resources assessment. *ISH J Hydraul Eng* 15(3):27–39
- Patra KC (2001) Hydrology and water resources engineering. Narosa Publishing House, New Delhi

Suribabu CR, Sudarsan JS, Ramanan SR (2015) A study on consecutive day maximum rainfall for disaster mitigation in Tiruchirapalli City, India. *Jordan J Civil Eng* 9(2):197–202

Vivekanandan N, Mathew FT (2010) Probabilistic modeling of annual d-day maximum rainfall. *ISH J Hydraul Eng* 16(1):122–133